# Effects of speech style, room acoustics, and vocal fatigue on vocal effort

Pasquale Bottalico,[a] Simone Graetzer, and Eric J. Hunter
*Voice Biomechanics and Acoustics Laboratory, Department of Communicative Sciences and Disorders, Michigan State University, East Lansing, Michigan 48824, USA*

(Received 6 October 2015; revised 12 February 2016; accepted 4 May 2016; published online 19 May 2016)

Vocal effort is a physiological measure that accounts for changes in voice production as vocal loading increases. It has been quantified in terms of sound pressure level (SPL). This study investigates how vocal effort is affected by speaking style, room acoustics, and short-term vocal fatigue. Twenty subjects were recorded while reading a text at normal and loud volumes in anechoic, semi-reverberant, and reverberant rooms in the presence of classroom babble noise. The acoustics in each environment were modified by creating a strong first reflection in the talker position. After each task, the subjects answered questions addressing their perception of the vocal effort, comfort, control, and clarity of their own voice. Variation in SPL for each subject was measured per task. It was found that SPL and self-reported effort increased in the loud style and decreased when the reflective panels were present and when reverberation time increased. Self-reported comfort and control decreased in the loud style, while self-reported clarity increased when panels were present. The lowest magnitude of vocal fatigue was experienced in the semi-reverberant room. The results indicate that early reflections may be used to reduce vocal effort without modifying reverberation time. 
[http://dx.doi.org/10.1121/1.4950812]

[CGC] Pages: 2870–2879

## I. INTRODUCTION

The interaction between a person, a room, and an activity leads to different sensations relating to voice production. This interaction determines acoustic comfort, which contributes to well-being. It also affects vocal comfort, which is a psychological measure that is determined by those aspects that reduce the vocal effort (Titze, 2000), e.g., the speaker-listener distance and the background noise level. Vocal comfort appears to decrease with the speaker's perceived fatigue and the sensation of needing to increase the voice level (Pelegrín-García and Brunskog, 2012). A speaker may unconsciously balance vocal effort and vocal comfort to maintain their own clarity and intelligibility.

Vocal control can be defined as the capacity to self-regulate vocal behavior. The sensation of control relates to the ability to adjust the voice consciously. In adverse conditions, speakers try to control their voice production in order to meet the needs of listeners (e.g., Wassink *et al.*, 2007; Hazan and Baker, 2011). For example, when conversing with a listener who has hearing limitations, a typical talker uses "clear speech." Such speech has been characterized by a slower speech rate, a wider range of fundamental frequency (fo), and a higher temporal modulation index than conversational speech (e.g., Picheny *et al.*, 1985; Ferguson *et al.*, 2010).

Vocal effort is a physiological measure that accounts for changes in voice production, which can be expressed by the A-weighted sound pressure level (SPL) (dB) at a 1 m distance from the mouth (ISO 9921, 2002). It relates to various factors such as the type of interlocutor (Hazan and Baker, 2011), the speaker-listener distance, the background noise level, and other acoustic characteristics of the room (Black, 1950; Pelegrín-García *et al.*, 2011), linguistic factors such as vowel quality (Eriksson and Traunmüller, 1999), and the speaker's level of fatigue (Rantala *et al.*, 2002; Laukkanen and Kankare, 2006).

Vocal fatigue is often experienced by speakers who use their voice for long periods and and/or with increased vocal effort, such as teachers. Titze (1999) identified two physiological aspects of such fatigue: laryngeal muscle fatigue and laryngeal tissue fatigue. Laryngeal muscle fatigue, which can involve tension in the vocal folds, appears to be caused by a depletion or accumulation of biochemical substances in the muscle fibers. Laryngeal tissue fatigue takes place in non-muscular tissue layers (epithelium, superficial, and intermediate layers of the lamina propria) and appears to be caused by temporary changes in molecular structure that result from mechanical loading and unloading (i.e., phonation; Titze, 2000). The minimization of vocal fatigue is particularly important when (1) the speaker is at high risk of vocal injury, such as in teaching environments (Hunter and Titze, 2010), when classroom acoustics are poor (Bottalico and Astolfi, 2012); and (2) when vocal function is impaired by loading and/or incomplete muscle recovery (Hunter and Titze, 2009).

SPL, in particular, has been found to be affected by vocal loading, possibly inducing vocal fatigue. Rantala *et al.* (2002) analyzed the recordings of 33 female teachers during the first and the last lessons (35–45 min) of a normal work

[a]Electronic mail: pb@msu.edu

 

day (5 h). The investigators divided the teachers into two categories: subjects who reported frequent symptoms of vocal fatigue (MC, many complaints), and subjects with few vocal complaints (FC, few complaints). SPL was found to increase by 0.5 dB between the first and last lesson, but this finding did not reach significance. Laukkanen and Kankare (2006), who examined male teachers' voices before and after a working day with the same division of subjects into groups, found that SPL increased in both groups.

The maximization of intelligibility, clarity, vocal comfort and control, and the minimization of vocal fatigue and effort, should be the priority of any professional talker. This is particularly important when the person is at elevated risk of vocal injury, such as in the case of teachers (Titze and Hunter, 2015). In classrooms, noise levels are typically high and acoustical conditions are not optimized for the talker but for the listener (Bottalico and Astolfi, 2012).

Reverberation time has been found to influence voice power level and vocal intensity in continuous speech. The effects on voice power level of reverberation time and speaker-listener distance were investigated by Pelegrín-García *et al.* (2011). Thirteen male talkers were recorded in four different environments: an anechoic chamber, a lecture hall, a corridor, and a reverberant room with average reverberation time ($T_{30,0.5\text{–}1\,\text{kHz}}$) of 0.04 s, 1.88 s, 2.34 s, and 5.38 s, respectively. The voice power level was found to depend almost linearly on the logarithm of the speaker-listener distance (with slopes between 1.3 and 2.2 dB per doubling of distance) and changed significantly among rooms (intercepts between 54.8 and 56.8 dB). With the exception of the reverberant room, voice power level decreased as reverberation time increased. Black (1950) reported an analysis of SPL measured in the context of read speech produced by 23 males in 8 rooms differing in shape (rectangular and drum), size (4.2 $m^3$ and 45.3 $m^3$), and reverberation time (0.2–0.3 s and 0.8–1.0 s). Greater vocal intensity was found in less reverberant room than in more reverberant rooms. Moreover, when comparing the first 3 and last 3 phrases of a total of 12 phrases, the speakers' mean relative intensity was found to be lower in more reverberant rooms than in less reverberant rooms.

Brunskog *et al.* (2009) investigated objectively measurable parameters of the rooms related to (1) any increase of the voice sound power, which is strongly correlated with vocal effort, produced by speakers, and (2) the speakers' subjective judgments about six different rooms with different sizes, reverberation times, and other physical attributes. The authors found that voice power is correlated with room size and inversely correlated with the magnitude of amplification by the room of the talker's voice at his/her ears, compared to anechoic conditions (termed "room gain" and "voice support").

In summary, previous research suggests that the speech level increases as vocal fatigue increases (Rantala *et al.*, 2002; Laukkanen and Kankare, 2006) and decreases under more reverberant conditions (Black, 1950; Pelegrín-García *et al.*, 2011). Unfortunately, higher reverberation times have been associated with decreases in speech intelligibility for students (Bradley, 1986; Astolfi *et al.*, 2012). Hence, it is necessary to study acoustical parameters that can be varied independently of the reverberation time and that can be used by speakers to decrease their vocal effort.

In this study, the effects of room acoustics, voice style (corresponding to normal and raised levels), and chronological task order or "experimental presentation order" on vocal effort (SPL) and self-reported vocal effort, control, comfort and clarity, are examined. Two independent room acoustic parameters were considered: reverberation time and external auditory feedback.

The main research questions of this study were (1) is it possible to decrease speakers' vocal effort by increasing their external auditory feedback, and (2) if there is such an effect, how does it interact with reverberation time and speech style effects?

## II. EXPERIMENTAL METHOD

The speech of 20 talkers was recorded in 3 different rooms in the presence of classroom babble, with and without reflective polycarbonate panels at 0.5 m from the talkers' mouths. The speech signals were processed to calculate measures of SPL.

### A. Subjects

Ethics approval for the experiment was granted by the Michigan State University Human Research Protection Program (IRB 13-1149). Twenty subjects, ten males and ten females, participated in the experiment. The subjects, who were non-smoking English-speaking university students, were aged between 18 and 30 years (mean age 20.8 years) and self-reported normal speech and hearing.

### B. Instructions and conditions

The subjects were instructed to read a text of about 30 s in duration in the presence of classroom babble noise, with and without reflective panels at 0.5 m from the mouth. The text was a six sentence excerpt from the Rainbow passage (Fairbanks, 1960) printed and attached to a music stand at 1 m from the subject. Two different speech styles were elicited: normal and loud. The instructions given for the styles were as follows: "Speak in your normal voice" (normal); "Imagine you are in a classroom and you want to be heard by all of the children" (loud).

Subjects were recorded in three different acoustic environments. The first was an anechoic room with dimensions 3.4 × 4.6 × 2.4 m. The second was a semi-reverberant room, 8.5 × 7.3 × 4.6 m. The walls were concrete block with some shelves and similar furniture covering some of the space, the ceiling was concrete, and the floor was covered with vinyl tile. The third room was a reverberant room with dimensions 7.7 × 6.4 × 3.6 m.

In each room, the subject was asked to read in 4 conditions (for a total of 12 tasks per subject): (i) with a normal vocal effort and without reflective panels; (ii) with a loud vocal effort and without reflective panels; (iii) with a normal vocal effort and with reflective panels; and (iv) with a loud vocal effort and with reflective panels. With the aim of an equal distribution of any (short-term) vocal fatigue across all the tasks, and to control for any unknown confounding variables relating to task order, the order of administration of the tasks was randomized. In sum, there were 12 tasks or conditions: 2 speech styles, two panel conditions, and 3 room environments.

Subjects answered four questions after each task: (i) Effort: How effortful was it to speak in this condition? (ii) Control: How well were you able to control your voice in this condition? (iii) Comfort: How comfortable was it to speak in this condition? (iv) Clarity: How clearly did you perceive your own voice in this condition? Subjects responded by making a vertical tick on a continuous horizontal line of 100 mm in length (on a visual analogue scale or VAS). The score was measured as the distance of the tick from the left end of the line. The extremes of the lines were "not at all" (left) and "extremely" (right).

## C. Equipment

Speech was recorded by a head-mounted omnidirectional microphone placed 5–7 cm from the mouth (Glottal Enterprises M80, Glottal Enterprises, Syracuse, NY). The microphone was connected to a personal computer (PC) via an external sound board (Scarlett 2i4 Focusrite, Focusrite, Windsor House, UK). The signals were recorded with Audacity 2.0.6 (SourceForge, La Jolla, CA) with a sampling rate of 44 100 Hz.

## D. Room acoustic parameters

Room acoustic parameters were obtained from the impulse response measurements in the non-occupied condition for the three rooms (ISO 3382–2, 2008). Balloon pops were used as impulses. The average reverberation times, $T_{30}$, for combined 500 Hz and 1 kHz octave bands, were determined for each room in four different positions. The $T_{30}$ for combined 500 Hz and 1 kHz octave bands was 0.04 s [standard deviation (s.d.) = 0.005] in the anechoic room, 0.78 s (s.d. = 0.012) in the semi-reverberant room and 2.37 s (s.d. = 0.167) in the reverberant room. The measured values of the reverberation time for the three rooms between 125 and 8 kHz are given in Table I.

To manipulate the level of external auditory feedback in the position of the talker, two reflective panels were placed at 45°, 0.5 m from the subject. The panels were made of transparent polycarbonate material and had a surface area of $56 \times 66\,\text{cm}^2$, which was perpendicular to the lines joining the panels and the subject. The presence of the panels generated a strong first reflection of the subject's voice. In order to quantify this effect, pink noise was emitted from the mouth and received by the ears of a Head and Torso Simulator with Mouth Simulator (HATS, 45BC KEMAR, G.R.A.S. Sound & Vibration, Holte, Denmark). This measurement was repeated in the three rooms, each room with and without reflective panels, maintaining a constant source (mouth) power. The ears were connected to an audio analyzer (XL2, NTI Audio, Schaan, Liechtenstein). Figure 1 shows the difference between the SPL measured per octave band in the anechoic room without panels and the sound levels measured in all room and panel conditions. A higher SPL was recorded in the frequencies relevant to speech in all rooms when panels were present. The higher the reverberation time in the room, the higher the increase in SPL introduced by the panels.

Babble noise was present during the entirety of the experiment. During all 12 tasks performed by each subject, classroom babble was emitted by a directional loudspeaker placed 2 m in front of the subject. The power level of the loudspeaker was set in order to obtain an A-weighed equivalent level averaging both ears of 62 dB at the talker position (measured with the HATS). This level represents the background noise present in a classroom during group activities (Shield and Dockrell, 2004).

## E. Processing of the voice recording

Analysis of the SPL was performed with MATLAB2015b. For each condition (task), a time history of A-weighted SPL in one-octave bands between 125 Hz and 4 kHz at 0.125 s intervals was obtained from the recorded speech. A correction factor accounting for the increase in SPL at the head-mounted microphone in the different combinations of room and panels was applied. The correction values are reported in Table II. The correction factors were measured by analyzing the SPL produced by the HATS, reproducing pink noise with a constant sound power level in the different conditions, at the head-mounted microphone, which was placed on the HATS.

After the correction of SPL, 12 time histories were obtained per subject. The average among all SPL values was computed per subject, and this mean was subtracted from each value of the 12 time histories performed by that subject. This within-subject centering was performed in order to evaluate the variation in the subject's vocal behavior in the different conditions from the "mean" vocal behavior. After transformation, the parameter was termed ΔSPL. The time information associated with the time histories (which

TABLE I. Reverberation time ($T_{30}$) and standard deviations measured in the three rooms per octave band. The babble noise spectrum (dBA) was measured per octave band using the HATS.

| Parameter (s.d.) | Room | 125 Hz | 250 Hz | 500 Hz | 1000 Hz | 2000 Hz | 4000 Hz | 8000 Hz |
|---|---|---|---|---|---|---|---|---|
| $T_{30}$ (s) | Anechoic | 0.05 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 |
| | | (0.005) | (0.007) | (0.006) | (0.005) | (0.005) | (0.005) | (0.007) |
| $T_{30}$ (s) | Semi-reverberant | 1.01 | 0.92 | 0.78 | 0.79 | 0.79 | 0.73 | 0.55 |
| | | (0.02) | (0.06) | (0.02) | (0.01) | (0.02) | (0.02) | (0.02) |
| $T_{30}$ (s) | Reverberant | 1.26 | 1.82 | 2.23 | 2.52 | 2.33 | 1.66 | 1.01 |
| | | (0.06) | (0.03) | (0.06) | (0.04) | (0.11) | (0.05) | (0.02) |
| Babble noise spectrum (dBA) | | 34.8 | 46.6 | 58.1 | 59.0 | 48.3 | 40.1 | 27.1 |

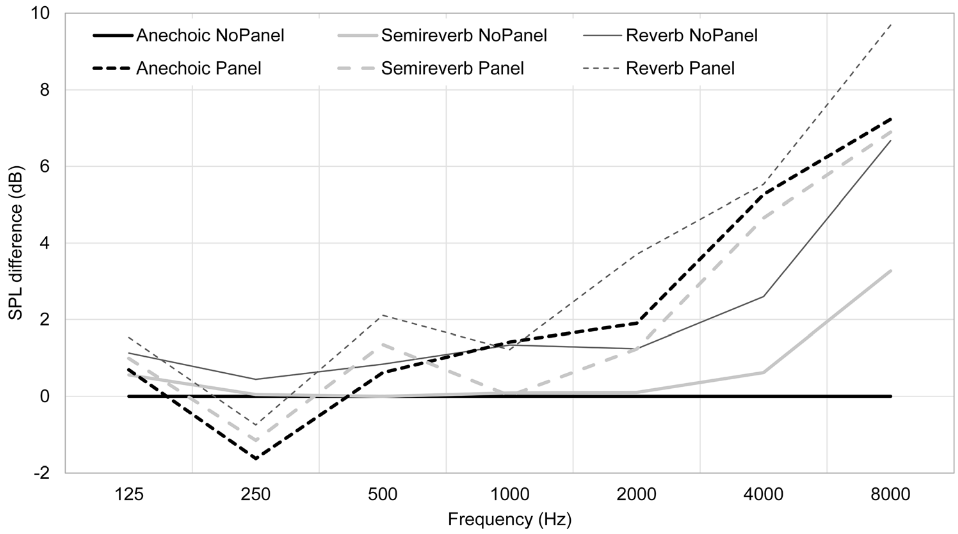


FIG. 1. Differences in SPL measured per octave band between the anechoic room without panels and the sound levels measured in all room and panel conditions.

typically ranged from 0 to 30 s within a task) was retained for inclusion in the statistical analysis.

### F. Statistical method

Statistical analysis was conducted using R version 3.1.2 (R Development Core Team, 2011). linear mixed-effects (LME) models were fit by restricted maximum likelihood (REML). Random effects terms were chosen on the basis of variance explained. Models were selected on the basis of the Akaike information criterion (Akaike, 1973; the model with the lowest value being preferred) and the results of likelihood ratio tests (a significant result indicating that the more complex of the two nested models in the comparison is preferred) and were built using lme4, lmerTest, and multcomp packages. Tukey's *post hoc* pairwise comparisons were performed to examine the differences between all levels of the fixed factors of interest. These are pairwise $z$ tests, where the $z$ statistic represents the difference between an observed statistic and its hypothesized population parameter in units of the standard deviation. The $p$ values for these tests were adjusted using the default single-step method (Hothorn *et al.*, 2008). The LME output includes the estimates of the fixed effects coefficients, the standard error associated with the estimate, the degrees of freedom, df, the test statistic, $t$, and the $p$ value. The Satterthwaite method (Satterthwaite, 1946) is used to approximate degrees of freedom and calculate $p$ values.

## III. RESULTS

### A. Vocal effort as SPL

A LME model was run with the response variable ΔSPL (dB) and the fixed factors (1) style, (2) room, (3) panel, and (4) chronological task order ("order") with interactions of (5) room and style, (6) room and order, (7) style and gender. The random effects were the interaction of subject and time (where time was measured in ms per task). Other possible interactions were excluded after likelihood-ratio tests indicated that their inclusion did not improve the model fit ($p > 0.1$). Model results are shown in Table III, while summary statistics in the 12 conditions are shown in Table IV. The effects of style, room, panel, and order, and the interactions between room and style, room and order, and style and gender were significant, with the exception of the interaction between order and the anechoic room.

The mean increase in ΔSPL from the normal to the loud style was 7.88 dB. However, the mean increase was greater in female subjects (8.64 dB) than in male subjects (7.09 dB). With regard to room, as shown in Fig. 2, the difference in ΔSPL between the styles was greater in the anechoic and reverberant rooms (8.15 and 7.96 dB, respectively) than in the semi-reverberant room (7.55 dB). In the normal style, the highest ΔSPL values were measured in the semi-reverberant room, and the lowest in the reverberant room; a higher ΔSPL in this style was measured in the semi-reverberant room (−3.63 dB) than the anechoic (−3.90 dB) and the reverberant

TABLE II. Increase in SPL in dB at the head-mounted microphone due to sound reflections, measured with a Head and Torso Simulator. The reference condition for SPL is the anechoic room without reflective panels.

| | | Frequency (Hz) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Room | Panel | 125 Hz | 250 Hz | 500 Hz | 1000 Hz | 2000 Hz | 4000 Hz | 8000 Hz |
| Anechoic | Absent | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | Present | 0.17 | −0.53 | 0.96 | 0.08 | 0.23 | 0.10 | 0.17 |
| Semi-reverberant | Absent | 0.14 | 0.04 | 0.13 | 0.10 | 0.10 | 0.17 | 0.14 |
| | Present | 0.14 | −0.31 | 0.97 | 0.17 | 0.07 | 0.47 | 0.14 |
| Reverberant | Absent | 0.01 | 0.10 | 0.20 | 0.27 | 0.20 | 0.23 | 0.01 |
| | Present | 0.21 | −0.28 | 0.88 | 0.30 | 0.40 | 0.37 | 0.21 |

TABLE III. A LME model fit by REML for the response variable ΔSPL (dB) including as fixed factors the terms (1) style, (2) room, (3) panel, and (4) chronological order ("order") with interactions of (5) room and style, (6) room and order, (7) style and gender, and the interaction of subject and time as a random effects term. As within-subject normalization was performed, it is not possible to appreciate whether or not there was a main effect of gender on ΔSPL. Reference levels are the normal style, the semi-reverberant room, absent panels, and female gender. R. = room; st. = style; Signif. = Significance; Semi-reverb = Semi-reverberant.

| Fixed factors | Estimate (dB) | Standard error (dB) | df | $t$ | $p$[a] |
|---|---|---|---|---|---|
| (Intercept) | −4.35 | 0.22 | 24164 | −20.01 | <0.001*** |
| Loud style | 8.33 | 0.19 | 34642 | 44.85 | <0.001*** |
| Anechoic room | −0.67 | 0.30 | 33596 | −2.23 | <0.05* |
| Reverberant room | −1.79 | 0.29 | 31671 | −6.25 | <0.001*** |
| Panel present | −0.81 | 0.09 | 34501 | −8.66 | <0.001*** |
| Order | 0.11 | 0.02 | 31927 | 4.35 | <0.001*** |
| Anechoic R.: Loud st. | 0.52 | 0.23 | 34470 | 2.28 | <0.05* |
| Reverberant R.: Loud style | 0.47 | 0.23 | 34526 | 2.05 | <0.05* |
| Anechoic R.: Order | 0.05 | 0.04 | 27015 | 1.36 | 0.174 |
| Reverberant R.: Order | 0.13 | 0.04 | 23743 | 3.58 | <0.001*** |
| Normal style: Male | 0.82 | 0.17 | 6671 | 4.87 | <0.001*** |
| Loud style: Male | −0.72 | 0.17 | 6536 | −4.31 | <0.001*** |

[a]Signif. codes: "***" < 0.001, "**" < 0.01, "*" < 0.05, "." < 0.1.

room (−4.58 dB). In the loud style, ΔSPL decreased as reverberation time increased, at 4.25, 3.92, and 3.38 dB for anechoic, semi-reverberant, and reverberant rooms, respectively ($T_{30}$ = 0.04, 0.78, and 2.37 s). *Post hoc* comparisons confirmed that, overall, the ΔSPL measured for the reverberant room was lower than both that of the semi-reverberant room ($z = -6.10$, $p < 0.001$) and the anechoic room ($z = -5.42$, $p < 0.001$), and that the difference between the anechoic and semi-reverberant rooms was not significant ($z = -0.62$, $p = 0.81$). Regarding the effect of panels, ΔSPL decreased when panels were present in both style conditions and in all three room conditions. The reduction in ΔSPL when panels were present rather than absent was 0.86 dB, as shown in Fig. 3.

The statistical model indicated a significant effect of chronological task order on ΔSPL and an interaction between order and room. The task order was randomized; consequently, each subject had a different task order. The effect of chronological task order on ΔSPL represents the compensations in a speaker's voice production over the time period of the recording session. It can be interpreted as an effect of short-term vocal fatigue.

In order to better understand the interaction between chronological task order and room, three simple linear regression models were fit to ΔSPL, one per room, with order as a predictor variable. The models that best fit the data in anechoic, semi-reverberant, and reverberant rooms are reported in Eqs. (1), (2), and (3), respectively,

$$\Delta SPL_{anechoic} = -1.24 + 0.22 \cdot \mathrm{Order}, \tag{1}$$

$$\Delta SPL_{semi-reverberant} = -0.53 + 0.12 \cdot \mathrm{Order}, \tag{2}$$

$$\Delta SPL_{reverberant} = -1.69 + 0.20 \cdot \mathrm{Order}, \tag{3}$$

where "Order" represents the chronological order of task administration from 1 to 12. The $p$ values associated with the factor of order in the three models were lower than 0.001. When compared with null models, the results of likelihood ratio tests were also significant at $p < 0.001$ in each case, confirming that the models including the Order term were preferred.

TABLE IV. Summary statistics for the variable ΔSPL (dB) in the 12 conditions (2 styles, 3 rooms, and 2 panels).

| Conditions | | | ΔSPL | | |
|---|---|---|---|---|---|
| Style | Room | Panel | Mean (dB) | Standard deviation (dB) | Standard error (dB) |
| Normal | Semi-reverberant | Absent | −3.16 | 8.79 | 0.16 |
| Normal | Semi-reverberant | Present | −4.10 | 8.76 | 0.16 |
| Normal | Anechoic | Absent | −3.50 | 8.61 | 0.16 |
| Normal | Anechoic | Present | −4.30 | 8.90 | 0.16 |
| Normal | Reverberant | Absent | −4.30 | 8.21 | 0.15 |
| Normal | Reverberant | Present | −4.85 | 8.12 | 0.14 |
| Loud | Semi-reverberant | Absent | 4.29 | 10.61 | 0.19 |
| Loud | Semi-reverberant | Present | 3.54 | 10.75 | 0.19 |
| Loud | Anechoic | Absent | 4.66 | 10.73 | 0.19 |
| Loud | Anechoic | Present | 3.84 | 10.79 | 0.19 |
| Loud | Reverberant | Absent | 3.99 | 9.81 | 0.17 |
| Loud | Reverberant | Present | 2.78 | 9.53 | 0.16 |

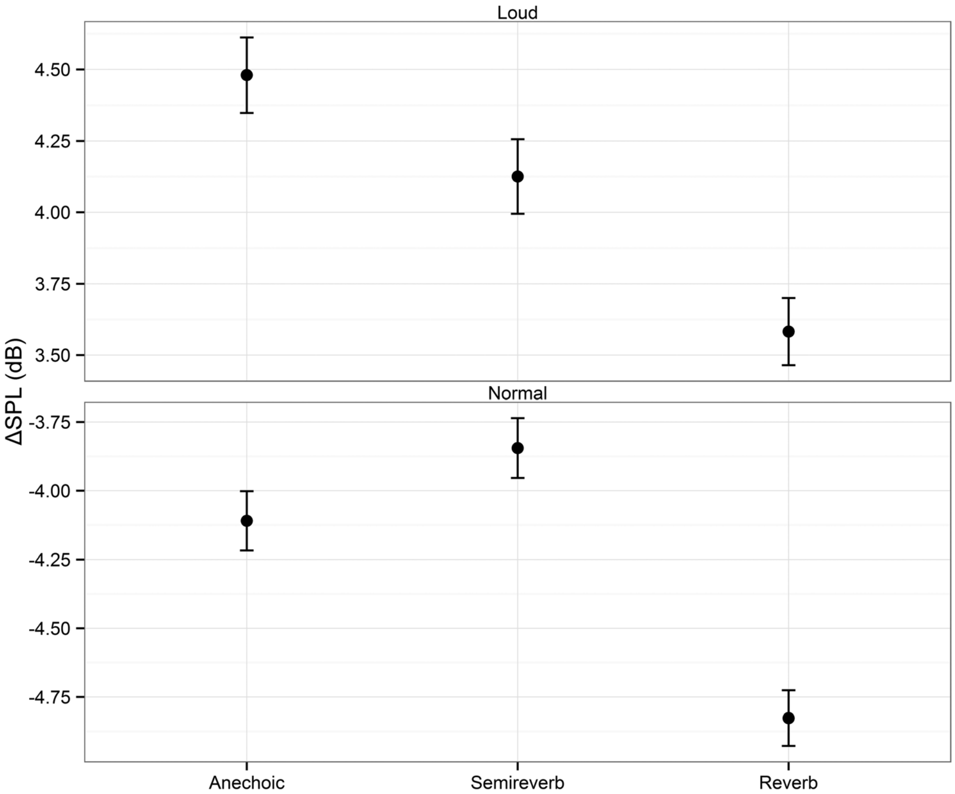


FIG. 2. Mean ΔSPL in dB across subjects per room for the loud (upper) and normal (lower) styles, where the error bands indicate ± standard error.

## B. Self-reported vocal effort, control, comfort, and clarity

Four separate LME models were run with the subjective response variables effort, control, comfort, and clarity, each with the fixed factors (1) style, (2) room, and (3) panel, and the random effects term of subject (Table V). Each response variable is reported in percent. The reference levels were the normal style, the anechoic room, and absent panels. Summary statistics for the self-reported variables are reported in Table VI.

The estimate for self-reported vocal effort in the loud style was 27.48% higher than that in the normal style. In the semi-reverberant and reverberant rooms, estimates were 5.11% and 9.89% lower, respectively, than the estimate associated with the anechoic room. The estimate for self-reported vocal effort in the presence of the panels was 4.95% lower than that without panels. These values are very similar to the actual differences in means. The effect of panels on self-reported effort is shown in Fig. 3. A Spearman's rho test indicated a significant relationship between self-reported effort and ΔSPL [$r_s(240) = 0.52, p < 0.001$].

The model estimate for self-reported vocal control was 10.28% lower in the loud style than in the normal style, while the estimate for self-reported vocal comfort was 20.35% lower in the loud style than in the normal style. The estimate for self-reported vocal clarity in the presence of the panels was 4.38% higher than that without panels ($p = 0.06$). These differences are again very similar to the actual difference in means. Other factors did not have observable effects.

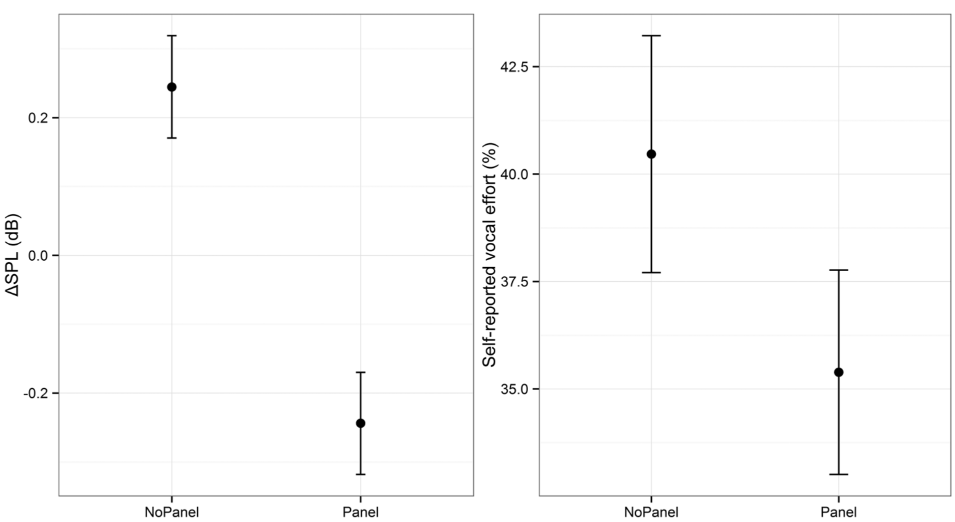


FIG. 3. (Left) Mean ΔSPL in dB across subjects per panel condition and (Right) self-reported vocal effort across subjects per panel condition, where error bands indicate ± standard error.

TABLE V. LME models fit by REML for the subjective response variables effort, control, comfort, and clarity, including fixed factors style, room, panel, and a random effects term: subject. Reference levels are normal style, anechoic room, and absent panels.

| Fixed factors | Estimate (—) | Standard error (—) | df | $t$ | $p$[a] |
|---|---|---|---|---|---|
| Effort | | | | | |
| (Intercept) | 31.60 | 4.39 | 40 | 7.19 | <0.001*** |
| Loud style | 27.48 | 2.46 | 215 | 11.13 | <0.001*** |
| Semi-reverberant R | −5.11 | 3.01 | 215 | −1.69 | 0.092. |
| Reverberant R | −9.89 | 3.02 | 215 | −3.27 | 0.001** |
| Panel present | −4.95 | 2.46 | 215 | −2.00 | <0.05 * |
| Control | | | | | |
| (Intercept) | 75.5 | 3.95 | 42 | 19.12 | <0.001*** |
| Loud style | −10.28 | 2.28 | 218 | −4.51 | <0.001*** |
| Semi-reverberant R | −0.74 | 2.8 | 215 | −0.27 | 0.790 |
| Reverberant R | 1.23 | 2.8 | 215 | 0.44 | 0.662 |
| Panel present | 1.13 | 2.29 | 215 | 0.5 | 0.621 |
| Comfort | | | | | |
| (Intercept) | 73.7 | 4.23 | 40 | 17.42 | <0.001*** |
| Loud style | −20.35 | 2.38 | 215 | −8.55 | <0.001*** |
| Semi-reverberant R | −0.96 | 2.91 | 215 | −0.33 | 0.741 |
| Reverberant R | 3.77 | 2.92 | 215 | 1.29 | 0.198 |
| Panel present | −0.22 | 2.38 | 215 | −0.09 | 0.927 |
| Clarity | | | | | |
| (Intercept) | 68.77 | 3.5 | 58 | 19.66 | <0.001*** |
| Loud style | 0.17 | 2.3 | 215 | 0.075 | 0.940 |
| Semi-reverberant R | −1.23 | 2.84 | 215 | −0.43 | 0.666 |
| Reverberant R | 0.78 | 2.85 | 215 | 0.275 | 0.783 |
| Panel present | 4.38 | 2.32 | 215 | 1.88 | 0.061. |

[a]Signif. codes: "***" < 0.001, "**" < 0.01, "*" < 0.05, "." < 0.1.

Figure 4 shows the mean self-reported vocal effort in the three rooms, for both normal and loud styles, with and without panels. It is apparent that the perception of vocal effort and reverberation time were inversely proportional and that the presence of panels was generally associated with a lower vocal effort. The only exception to this rule was the condition with the loud style in the reverberant room. This exception may be due to excessive energy in the reflections because of the combination of the reverberant sound field and the increased first reflection associated with the panels.

## IV. DISCUSSION

In this study, the effects were measured of speech style, reverberation time, panel, and chronological task order on ΔSPL and on four subjective variables relating to voice production: effort, control, comfort, and clarity. With regard to speech level, a difference in ΔSPL of 7.88 dB was found between the normal and loud styles. In the semi-reverberant room, the loud style was associated with an increase of 7.5 dB relative to the normal style. The increase was ∼8 dB for the anechoic and reverberant rooms. The standard ISO 9921 (2002) indicates a variation of 6 dB between steps of increasing volume in speech. However, the descriptors associated with vocal effort in the standard are normal, raised, and loud, and in the present study, the loud style corresponds to some point in between the raised and loud levels of the standard in terms of SPL. This difference could be explained on the basis of the number of instructions about voice style: fewer instructions will correspond to a larger dynamic range of the voice across styles. Bottalico *et al.* (2015), in a similar experiment that involved instructing talkers to use three voice styles (soft, normal, and loud), found a smaller dynamic range than the one reported in the present study (6.8 dB between loud and normal, and 7.8 dB between normal and soft). A second possible explanation could be that the dynamic range of speech in noisy conditions might tend to be smaller than the range in quiet conditions (Bottalico *et al.*, 2015). This might occur because speech produced in a noisy environment will be higher in level than speech produced in a quiet environment (Lombard effect), all else being equal. In noisy conditions, in the normal style, a speaker has a large dynamic range (both softer and louder) available and can adjust their level relatively freely. However, in noisy conditions, in the loud style, a speaker has a reduced range because a volume "saturation" point is likely to appear.

Across panel and style conditions, ΔSPL was found to be higher in the anechoic and semi-reverberant rooms than in the reverberant room. Consistent with the findings of Black (1950) and Pelegrín-García *et al.* (2011), in this study, as reverberation time increased (between 0.04 and 2.4 s), mean ΔSPL decreased (anechoic: 0.25, semi-reverberant: 0.24, reverberant: −0.47). The interaction between room and

TABLE VI. Summary statistics for the variables effort, control, comfort, and clarity in the 12 conditions (2 styles, 3 rooms, and 2 panels).

| Conditions | | | Effort (%) | | Control (%) | | Comfort (%) | | Clarity (%) | |
|---|---|---|---|---|---|---|---|---|---|---|
| Style | Room | Panel | Mean | Standard error | Mean | Standard error | Mean | Standard error | Mean | Standard error |
| Normal | Anechoic | Absent | 28.8 | 6.0 | 78.2 | 5.0 | 76.1 | 4.9 | 75.1 | 3.7 |
| Normal | Anechoic | Present | 24.0 | 4.5 | 75.8 | 4.9 | 72.1 | 4.9 | 72.9 | 4.0 |
| Normal | Semi-reverberant | Absent | 25.7 | 5.7 | 73.7 | 6.0 | 72.7 | 5.6 | 66.1 | 5.0 |
| Normal | Semi-reverberant | Present | 19.8 | 4.5 | 72.5 | 6.2 | 75.3 | 4.5 | 71.5 | 4.1 |
| Normal | Reverberant | Absent | 26.4 | 5.0 | 77.6 | 4.3 | 73.5 | 4.2 | 63.5 | 5.1 |
| Normal | Reverberant | Present | 20.0 | 4.6 | 79.8 | 3.3 | 77.9 | 3.3 | 75.6 | 3.6 |
| Loud | Anechoic | Absent | 65.2 | 6.6 | 62.8 | 5.6 | 53.4 | 7.0 | 65.5 | 5.4 |
| Loud | Anechoic | Present | 53.5 | 5.4 | 66.8 | 4.6 | 52.2 | 5.6 | 70.6 | 4.2 |
| Loud | Semi-reverberant | Absent | 55.9 | 6.1 | 64.7 | 5.0 | 49.2 | 6.1 | 66.2 | 5.0 |
| Loud | Semi-reverberant | Present | 49.6 | 5.3 | 69.7 | 4.7 | 52.6 | 5.2 | 75.4 | 3.8 |
| Loud | Reverberant | Absent | 40.1 | 6.1 | 66.4 | 4.6 | 62.5 | 5.3 | 75.7 | 5.3 |
| Loud | Reverberant | Present | 45.4 | 5.4 | 65.1 | 5.1 | 55.3 | 5.7 | 72.4 | 5.8 |

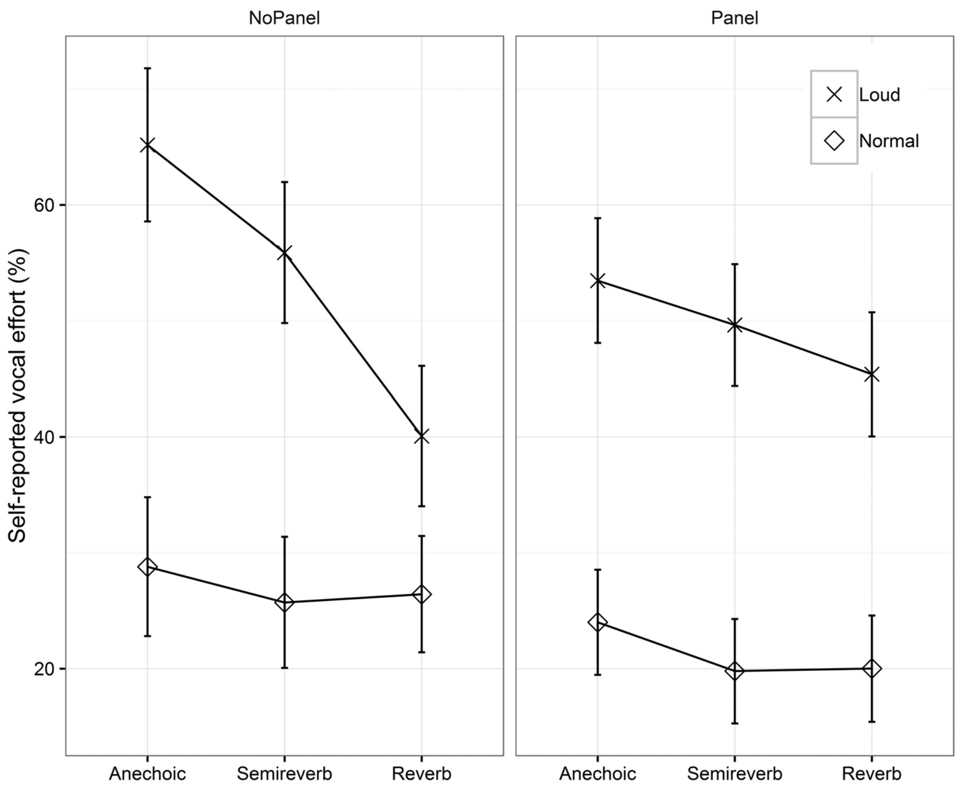


FIG. 4. Mean self-reported vocal effort in percent across subjects in the three rooms (anechoic, semi-reverberant, and reverberant) for normal and loud styles, with and without panels. Error bands indicate ± standard error.

style principally related to the relationship between the anechoic and semi-reverberant rooms in the two styles: ΔSPL was higher in the anechoic room than the semi-reverberant room in the loud style but lower in the normal style. In the loud style, the voice intensity was higher and, consequently, the reflected sound was more intense. In particular, when the energy emitted by the subject was higher, the reflections associated with the panels seem to be more effective in decreasing the subject's ΔSPL in the semi-reverberant room. In the loud style, ΔSPL was lower in the semi-reverberant room than in the anechoic room (mean difference = 0.34 dB). In the normal style, ΔSPL decreased from −4.58 dB in the reverberant room to −3.90 dB in the anechoic room and −3.63 in the semi-reverberant room, the semi-reverberant room being the smallest.

The effect of the panel being present was a decrease in the ΔSPL (mean = 0.86 dB), which was observable in all room and style conditions. The placement of the panels near the talker increased the reflected energy (and external auditory feedback) in the talker position, thus, increasing the levels of voice support and room gain, as defined by Pelegrín-García (2011). It is consistent with his findings that there was an inverse relationship between SPL and the quantity of reflected energy. In this study, as expected, when panels were present, talkers reported greater clarity of their own voice.

As within-subject normalization was performed, it is not possible to appreciate whether or not there was a main effect of gender on vocal effort. However, the interaction between gender and voice style was observed to be significant. Female subjects showed a larger dynamic range in the voice level between styles (i.e., mean difference between the styles) than male subjects. The dynamic range was 8.64 dB for females and 7.09 dB for males. A similar result was found by Bottalico *et al.* (2015), where the mean differences between the normal and loud styles were 7.50 dB for females and 6.32 dB for males in a typical classroom. The larger dynamic ranges associated with females could provide insights into reported gender-associated vocal health risks (e.g., Vilkman *et al.*, 1999; Titze *et al.*, 2003; Hunter and Titze, 2010; Hunter *et al.*, 2011).

An increase in ΔSPL across the 12 tasks was observed, which may indicate short-term vocal adjustment, likely a form of short-term fatigue. This finding is consistent with the tendency for SPL to increase with vocal loading observed by Rantala *et al.* (2002) and Laukkanen and Kankare (2006). Overall, reverberation time and SPL were inversely related such that as reverberation time increased from 0.78 s (semi-reverberant room) to 2.37 s (reverberant room), there was a decrease in ΔSPL of 0.72 dB. As expected, the same relationship was found between reverberation time and self-reported effort. In order to better interpret these findings, it was necessary to consider the interaction between reverberation time and order. The relationship between ΔSPL and short-term vocal adjustment or fatigue (evaluated by means of the chronological order from 1 to 12) was observed to strongly depend on room.

As reported in Eqs. (1)–(3), the slopes were 0.22 dB/Order in the anechoic room, 0.12 dB/Order in the semi-reverberant room and 0.20 dB/Order in the reverberant room. These values could be representative of the different effects of room on (short-term) vocal fatigue; they suggest that lower vocal demands and lower magnitudes of vocal fatigue were experienced by talkers in the room in which the reverberation time was more likely to be found in a typical space (the semi-reverberant room). Arguably, this hypothesis is confirmed by the results of Bottalico and Astolfi (2012). In fact, the authors stated that a range of mid-frequency reverberation time of between 0.75 and 0.85 s could be an optimal range for a talker in a classroom as it offers good voice support.

## V. CONCLUSIONS

In this study, 20 subjects performed vocal tasks in the presence of classroom babble noise in anechoic, semi-reverberant, and reverberant environments. In each of these environments, the room acoustics were modified by placing two reflective panels at 0.5 m from the subject. Normal and loud styles were elicited by instruction. After each task, the subject responded to questions addressing their perception of vocal effort, comfort, control, and the clarity of their own voice.

It has been demonstrated that the placement of reflective surfaces can improve the quality of the sound field for speakers. The increase in the external auditory feedback due to reflective panels significantly reduced vocal effort. SPL was observed to decrease by 0.86 dB when panels were present. That is to say, the subjects benefited in an objectively measurable way from the panels, and this benefit was also perceived by the subjects. While the effect of the panels was consistent among styles and rooms, the effect was strongest in the reverberant room (−0.92 dB), followed by the semi-reverberant room (−0.86 dB) and the anechoic room (−0.51 dB). Additionally, the effect of panels was stronger in the loud style (−0.93 dB), then in the normal style (−0.76 dB).

As far as the subjective evaluation of vocal effort is concerned, panels were generally associated with a lower perceived vocal effort, with the exception of the loud style and reverberant room condition. This may be due to excessive energy in the reflections due to the combination of the high reverberation time, the increase in the first reflection associated with the panels, and the higher speech level.

Previous research suggests that the speech level decreases under more reverberant conditions. This result was confirmed for the loud style but not for the normal style. Moreover, it was found that in rooms with different reverberation times there is a different rate of increase in the vocal effort (which is generally associated with vocal fatigue). These rates were 0.12 dB/task across the 12 tasks for the semi-reverberant room, 0.22 dB/task for the anechoic room, and 0.20 dB/task for the reverberant room.

This preliminary study of the differences in SPL and self-reported effort induced by changes in style and room acoustics confirms the sensitivity of vocal effort to the magnitude of auditory feedback. In this study, the reverberation times in two of the three rooms (anechoic and reverberant) were unusual in typical rooms. They were selected in order to cover the widest available range of reverberation time. In the future, experiments could be conducted in more typical environments, such as classrooms. Furthermore, in this study the reflective panels were placed at a single distance from the speaker. In order to improve classroom design and to be able to give recommendations concerning the placement of reflective surfaces, it is necessary to test the effect of panels on speech at different distances from the speaker and at different angles. Finally, for a more systematic evaluation of the effects of reflective panels, it will be necessary to perform some experiments in virtual acoustic environments for improved control of the acoustical parameters.

## ACKNOWLEDGMENTS

The authors would like to thank, for their assistance, the members of the Voice Biomechanics and Acoustics Laboratory, Michigan State University. Additionally, they would like to express their gratitude to the subjects involved in the experiment. The authors would also like to acknowledge the efforts of the Associate Editor and careful reviewers whose comments were invaluable to this work. This research was funded by the National Institute on Deafness and other Communication Disorders of the National Institutes of Health under Award No. R01DC012315. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Akaike, H. (**1973**). "Information theory and an extension of maximum likelihood principle," in *Second International Symposium on Information Theory*, edited by B. N. Petrov and F. Csaki (Akademiai Kiado, Budapest), pp. 167–281.

Astolfi, A., Bottalico, P., and Barbato, G. (**2012**). "Subjective and objective speech intelligibility investigations in primary school classrooms," J. Acoust. Soc. Am. **131**(1), 247–257.

Black, J. W. (**1950**). "The effect of room characteristic upon vocal intensity and rate," J. Acoust. Soc. Am. **22**(2), 174–176.

Bottalico, P., and Astolfi, A. (**2012**). "Investigations into vocal doses and parameters pertaining to primary school teachers in classrooms," J. Acoust. Soc. Am. **131**(4), 2817–2827.

Bottalico, P., Graetzer, S., and Hunter, E. J. (**2015**). "Effects of voice style, noise level, and acoustic feedback on objective and subjective voice evaluations," J. Acoust. Soc. Am. **138**(6), EL498–EL503.

Bradley, J. S. (**1986**). "Speech intelligibility studies in classrooms," J. Acoust. Soc. Am. **80**(3), 846–854.

Brunskog, J., Gade, A. C., Bellester, G. P., and Calbo, L. R. (**2009**). "Increase in voice level and speaker comfort in lecture rooms," J. Acoust. Soc. Am. **125**(4), 2072–2082.

Eriksson, A., and Traunmüller, H. (**1999**). "Perception of vocal effort and speaker distance on the basis of vowel utterances," in *Proc. International Conference on the Phonetic Sciences*, San Francisco, CA.

Fairbanks, G. (**1960**). "The rainbow passage," in *Voice and Articulation Drillbook*, 2nd. ed. (Harper and Brothers, New York).

Ferguson, S. H., Poore, M. A., Shrivastav, R., Kendrick, A., McGinnis, M., and Perigoe, C. (**2010**). "Acoustic correlates of reported clear speech strategies," J. Acad. Rehabil. Audiol. **43**, 45–64.

Hazan, V., and Baker, R. (**2011**). "Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions," J. Acoust. Soc. Am. **130**(4), 2139–2152.

Hothorn, T., Bretz, F., and Westfall, P. (**2008**). "Simultaneous inference in general parametric models," Biometrical J. **50**(3), 346–363.

Hunter, E. J., Tanner, K., and Smith, M. E. (**2011**). "Gender differences affecting vocal health of women in vocally demanding careers," Log. Phon. Voc. **36**, 128–136.

Hunter, E. J., and Titze, I. R. (**2009**). "Quantifying vocal fatigue recovery: Dynamic vocal recovery trajectories after a vocal loading exercise," Ann. Otol. Rhinol. Laryngol. **118**(6), 449–460.

Hunter, E. J., and Titze, I. R. (**2010**). "Variations in intensity, fundamental frequency, and voicing for teachers in occupational versus nonoccupational settings," J. Sp. Lang. Hear. Res. **53**, 862–875.

ISO (**2002**). ISO 9921:2002(E), "Ergonomics—Assessment of speech communication" (International Organization for Standardization, Geneva, Switzerland).

ISO (**2008**). ISO 3382-2:2008(E), "Acoustics–Measurement of room acoustic parameters, Part 2: Reverberation time in ordinary rooms" (International Organization for Standardization, Geneva, Switzerland).

Laukkanen, A., and Kankare, E. (**2006**). "Vocal loading-related changes in male teachers' voice investigated before and after a working day," Folia Phoniatr. Log. **58**, 229–239 (2006).

Pelegrín-García, D., and Brunskog, J. (**2012**). “Speakers’ comfort and voice level variation in classrooms: Laboratory research,” J. Acoust. Soc. Am. **132**, 249–260.
Pelegrín-García, D., Smits, B., Brunskog, J., and Jeong, C. (**2011**). “Vocal effort with changing talker-to-listener distance in different acoustic environments,” J. Acoust. Soc. Am. **129**(4), 1981–1990.
Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1985**). “Speaking clearly for the hard of hearing intelligibility differences between clear and conversational speech,” J. Speech Lang. Hear. Res. **28**(1), 96–103.
Rantala, L., Vilkman, E., and Bloigu, R. (**2002**). “Voice changes during working: Subjective complaints and objective measurements for female primary and secondary schoolteachers,” J. Voice **16**(4), 344–355.
R Development Core Team (**2011**). “R: A language and environment for statistical computing” (R Foundation for Statistical Computing, Vienna), available at http://www.R-project.org (Last viewed January 19, 2014).
Satterthwaite, F. E. (**1946**). “An approximate distribution of estimates of variance components,” Biometr. Bull. **2**(6), 110–114.
Shield, B., and Dockrell, J. (**2004**). “External and internal noise surveys of London primary schools,” J. Acoust. Soc. Am. **115**(2), 730–738.
Titze, I. R. (**1999**). “Toward occupational safety criteria for vocalization,” Log. Phon. Vocol. **24**, 49–54.
Titze, I. R. (**2000**). *Principles of Voice Production*. Second printing (National Center for Voice and Speech, Iowa City, IA), pp. 229–233, 361–366.
Titze, I. R., and Hunter, E. J. (**2015**). “Comparison of vocal vibration-dose measures for potential-damage risk criteria,” J. Speech, Lang. Hear. Res. **58**(5), 1425–1439.
Titze, I. R., Švec, J. G., and Popolo, P. S. (**2003**). “Vocal dose measures: Quantifying accumulated vibration exposure in vocal fold tissues,” J. Speech, Lang. Hear. Res. **46**(6), 919–932.
Vilkman, E., Lauri, E-R., Alku, P., Sala, E., and Sihvo, M. (**1999**). “Effects of prolonged oral reading on F0, SPL, subglottal pressure and amplitude characteristics of glottal flow waveforms,” J. Voice **13**(2), 303–312.
Wassink, A. B., Wright, R. A., and Franklin, A. D. (**2007**). “Intraspeaker variability in vowel production: An investigation of motherese, hyperspeech, and Lombard speech in Jamaican speakers,” J. Phon. **35**, 363–379.